

PC-Based Software for Analyzing U.S. Mining Accidents, Injuries, and Illnesses

Audrey F. Glowacki, Programmer/Analyst

Mining Health and Safety Research, Pittsburgh Research Center

Pittsburgh, Pennsylvania

Abstract. Since 1978, the Mine Safety and Health Administration (MSHA) has compiled two databases that track company and accident/ injury/illness information reported by U.S. mine operators and independent contractors. These databases are a valuable resource for both government and industry in their efforts to improve mine safety. However, the databases are large and complex, growing at the rate of twelve new files per year. Searching them without custom software can be tedious. Furthermore, because much of the data are stored in a coded format, paper manuals must be referenced when searching the data and interpreting search results.

To make the MSHA databases more easily accessible to the mining industry, the U. S. Bureau of Mines (USBM) has developed the Mine Accident Decision Support System (MADSS). MADSS is a DOS-based software application designed specifically for searching and analyzing the MSHA data. It allows users to quickly build and execute custom searches, and automates the process of linking and extracting data across multiple files. An online dictionary explains the MSHA data and the coding systems that are used to store information. MADSS can also convert coded data to their English descriptions before presenting the results of a search. Several other important features of MADSS are its ability to identify accident trends via its grouping and statistical capabilities, and to export search results to various DOS file formats. This export feature lets users easily transfer data to other software for advanced statistical analysis, report formatting, or graph generation.

MADSS and the MSHA data will be available in mid-1996. The software and data files will be distributed via the Internet and CD-ROM. While the software itself must be installed on the end user's PC, it will be possible to store the data on a shared network file server, or a local or shared CD-ROM.

THE MSHA DATA

MSHA compiles two databases that are useful for analyzing mining accidents. The first tracks "company" information, such as address, employment, and productivity data, for U.S. mine operators and independent contractors. For mine operators, details about the mine, such as mine name and type, are also documented. The second database tracks accidents, injuries, and illnesses reported by those companies (for brevity, this database will be referred to as the *Accident* database). With access to both databases, one can extract accidents for specific companies, or identify companies that experienced particular types of accidents.

However, the MSHA data files are not easy to search without special purpose software. The structure of the databases is complex, consisting of twelve files per year (figure 1). Three types of files are used to store company, accident, and accident narrative data. There are four files of each type, as MSHA stores data for different company types (operator or contractor) and commodity types (coal or metal/nonmetal) in separate files. So, storing five years of data requires maintaining sixty files. Unless the search process is automated, it can be tedious to search data for multiple years, company types, or commodity types, or extract a combination of company, accident, and narrative data. With MADSS, users simply identify the years, commodities, and company types of interest. The software handles the details of locating relevant files, linking data across files, and extracting requested data. The retrieved information is combined into a single table of results.

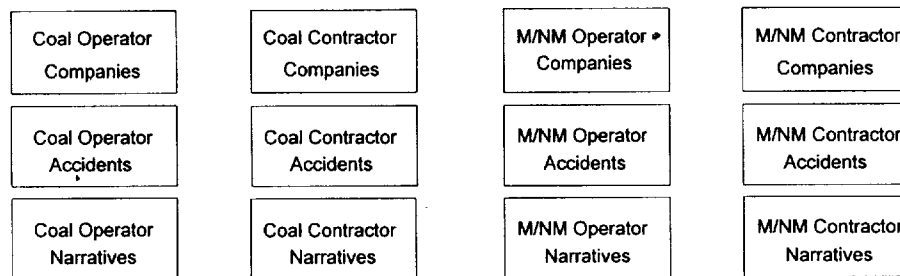


Figure 1: One year of MSHA data consists of twelve data files

The MSHA data can be confusing to a new user. Much of it is stored in a coded format. For example, the accident/injury/illness type for each accident is stored as a numeric code. With MADSS, pop up windows that explain each code, such as the one in figure 2, can be accessed at the press of a key. At the user's request, MADSS will convert coded data retrieved in a search to their equivalent English descriptions, producing results that are easy to interpret. MADSS also has a built-in data dictionary that gives detailed information about each MSHA data field.

CODE	ACC/INJ/ILL VALUES
	DESCRIPTION
00	NEC (CATEGORY DOES NOT INCLUDE OCCUPATIONAL ILLNESS)
01	DISORDERS/PHYSICAL AGENTS OTHER THAN TOXIC MATERIALS
02	DISORDERS/REPEATED TRAUMA
03	DUST DISEASE OF THE LUNGS
04	ELECTRICAL
05	ENTRAPMENT
06	EXPLODING VESSELS UNDER PRESSURE
07	EXPLOSIVES AND BREAKING AGENTS
08	FALL OF FACE, RIB, SIDE, OR HIGHWALL
09	FALL OF ROOF/FALL OF BACK
10	FALLING/ROLLING/SLIDING ROCK OR MATERIAL OF ANY KIND

Figure 2: Coding system for the field ACC/INJ/ILL

OPERATOR V.S. CONTRACTOR SEARCHES

Both mine operators and contractors are categorized as "companies" in MADSS. However, each submits company data to MSHA differently. An operator represents a single mine. A contractor, on the other hand, employs individuals who may work at more than one operator-owned mine. Both operators and contractors report general company information (such as owning company, address, employment, and production) to MSHA. Operators also include details specific to their mines, such as mine name and mine type. Mine-specific data are not reported by contractors, since contractors may have multiple mining operations.

As far as accidents are concerned, operators and contractors report data for their own employees, but report identical information for those incidents. The accident reports identify both the company that reported the incident, and the operator of the mine where the accident occurred. Consequently, accidents may be searched either by reporting company or mine of occurrence. Some of the other data that are tracked include victim age, sex, and work experience, type of accident, date of accident, and degree and source of injury.

To accommodate the reporting differences between operators and contractors, and to provide flexibility in searching the MSHA data, MADSS provides a choice of five Search Options (figure 3). This Search Option determines whether accidents will be searched by reporting company or mine of occurrence, and also whether operator or contractor data will be included. For example, when searching by mine of occurrence, the user can choose to include only operator accidents, only contractor accidents, or both operator and contractor accidents that occurred in the selected mines.

SEARCH OPTIONS

SEARCH REPORTING COMPANIES AND/OR THEIR ACCIDENTS

1. Operators <-> Operator accidents

2. Contracrors <-> Contractor accidents

SEARCH MINES AND/OR THEIR ACCIDENTS:

3. Operators <-> Operator accidents

4. Operators <-> Contractor accidents

5. Operators <-> Both operator and contractor accidents

Choose which companies and/or accidents to search (1-5) :

HELP
OK
CANCEL

Choose a company type for this search (OK=accept)

Figure 3: Available Search Options

CUSTOMIZING THE SEARCH

Once a Search Option is selected, additional constraints can be placed on the search. At minimum, these constraints must choose the years and commodity types (coal or metal/nonmetal) of interest for the search. This information, combined with the selected Search Option, determines which MSHA files will be searched. Optionally, additional search criteria may be defined to further constrain the companies and/or accidents to include from those files.

Search constraints are specified by building *Search Conditions* for relevant data fields in the MSHA databases. In each case, those Search Conditions identify a set of acceptable values for the selected field. Only records that have one of those values will be included in the search.

For example, to limit the years of the search, select the YEAR data field. A Search Conditions table for YEAR will appear on the screen (figure 4). This table is used to specify mathematical conditions that identify the acceptable values for the field. Here, one range condition is defined that selects all years greater than or equal to 1990, and less than or equal to 1993 (i.e., years 1990 through 1993). This condition uses the \geq (greater than or equal to) and \leq (less than or equal to) relational operators. However, other types of conditions can be constructed using any of the operators in Table 1. Up to ten Search Conditions can be listed per field - records that match at least one Search Condition for every field selected will be included in the search. For example, adding the Search Condition " $= 1994$ " to figure 4 would cause MADSS to search the MSHA files for the years 1990 through 1994.

YEAR			
OPERATOR	VALUE	OPERATOR	VALUE
\geq	1990	\leq	1993

Figure 4: Selecting the years 1990-1993

By default, all records for the selected years, commodity types and Search Option will be included in the search. However, one can constrain the search further by defining Search Conditions for appropriate Company and Accident fields.

Figure 5 shows a partial list of data fields that are tracked in the Company database (only the first page of this two-page list is shown). Search Conditions can be defined for any of these fields. For example, to select companies located in particular states, select the STATE OF LOC field and then enter Search Conditions to select those states. Figure 6 shows a Search Condition that limits the search to West Virginia.

Help Accept

ACCIDENT CONDITIONS - Page #1 of 3					
VICTIM					
___ Age	___ Regular Job Title	___ This Mine Exp			
___ Sex	___ Total Mining Exp	___ This Job Exp			
INJURY					
___ Acc/Inj/Ill	___ Part of Body	___ Document No.			
___ Accident Type	___ Nature of Injury	___ Closed Case Doc No			
	___ Closed Out				
TIME/LOCATION					
___ Accident Subunit	___ Date of Accident	___ Hours Into Shift			
___ Location	___ Shift Start Time	___ Time of Accident			

<Enter> Select | Delete | <Esc> Cancel | <PgDn> Nxt pg |
Select a data field for specifying Search Conditions

Figure 7: Partial list of accident data tracked by MSHA

ACC/INJ/ILL			
OPERATOR	VALUE	OPERATOR	VALUE
=	02		

Figure 8: Selecting injuries categorized as “disorders/repeated trauma”

Search Conditions may be defined for both Company and Accident fields, but they will be interpreted differently depending on whether the search extracts company records or accident records. When extracting company records, MADSS will retrieve companies that match the specified company criteria *and* are associated with accidents that match the specified accident criteria. When extracting accident records, MADSS will retrieve accidents that match the specified accident criteria *and* are associated with companies that match the specified company criteria.

CHOOSING DATA TO OUTPUT

Once the user has chosen which companies and/or accidents to include, he or she must specify whether the search will extract company or accident records, and pick the information to report for those records. This is done by selecting data fields from lists that are similar to those in figures 5 and 7. These data fields are referred to as the Search Output. If any Accident fields are selected, the search will extract accident records. Otherwise, the search will extract company records.

Figure 9 shows the Search Overview screen. This screen gives an overview of the search as it is being constructed. The leftmost table shows the fields for which Search Conditions have been defined, while the rightmost table shows the data fields to be included in the Search Output.

A sample search has been loaded. According to this figure, the Search Conditions limit the search to particular years, commodities, accident/injury/illness types, and injuries to particular parts of the body. The actual Search Conditions for these fields are not visible from this screen. However, they select the years 1990 through 1993, the coal and metal/nonmetal commodities, the accident/injury/illness type “disorders/repeated trauma,” and the part of body “wrist”.

Table 1: Valid mathematical operators for building Search Conditions

Operator	Description	Example
=	Matches all values equal to the specified value	= 1992
>	Matches all values greater than the specified value	> 1992
<	Matches all values less than the specified value	< 1992
>=	Matches all values greater than or equal to the specified value	>= 1992
<=	Matches all values less than or equal to the specified value	<= 1992
NOT	Matches all values except the specified value	NOT 1992
BLANK	Matches all blank values (note that no value is specified)	BLANK
NOT BLANK	Matches all nonblank values (note that no value is specified)	NOT BLANK
SUBSTR	Matches all values that contain the specified substring somewhere in the field (for searching accident narratives only) The example would match narratives such as "The victim was injured by a roof fall..."	SUBSTR ROOF

Help Accept

COMPANY CONDITIONS - Page #1 of 2

DATA REPORTED BY MINE OPERATORS

Mine Id

Mine Name

State of Loc

Inspection Office

Mine Type

County of Loc

SIC

Canvass Class

Seam Height

DATA REPORTED BY CONTRACTORS

Contractor ID

DATA REPORTED BY BOTH

CO Status

Owning Company

Mail State

Prior Status

Mail Address

Mail Zip Code

Status Change Date

Mail City

<Enter> Select | Delete | <Esc> Cancel | <PgDn> Nxt pg |
Select a data field for specifying Search Conditions

Figure 5: Partial list of Company data tracked by MSHA

STATE OF LOC			
OPERATOR	VALUE	OPERATOR	VALUE
=	WV		

Figure 6: Selecting companies located in West Virginia

Similarly, Search Conditions can be defined for fields in the Accident database to restrict the search to specific accidents, injuries, or illnesses. Figure 7 shows the first page of available Accident fields. By defining Search Conditions for appropriate data fields in this list, one can select accidents involving victims of particular ages, job titles, and/or work experience. Or, accidents can be selected based on accident type or the part of body that was injured. For example, figure 8 shows a Search Condition that selects injuries categorized with an accident/injury/illness type of "02". This code represents the value "disorders/repeated trauma" (see figure 2).

Help Conditions Output Empty Setup Go eXit

SEARCH OVERVIEW

Grouping OFF
Level 01

Search Conditions		Search Conditions	
Field Name	Op	Field Name	Format
YEAR	AND	YEAR	VALUE
COMMODITY	AND	DATE OF ACCIDENT	VALUE
PART OF BODY	AND	REGULAR JOB TITLE	DESC
ACC/ INJ/ ILL			

<Enter> Select | Delete | <F4> Nxt Tbl | <PgUp> Prv pg | <PgDn>Nxt Pg |
 Define Conditions and Output for this search - GO runs the search

Figure 9: The Search Overview screen

The Search Output fields selected for this search include YEAR, DATE OF ACCIDENT, and REGULAR JOB TITLE. Because the latter two fields reside in the Accident database, this search will extract accident records. Assuming that Search Option #5 was chosen, the search will extract a list of repeated trauma accident involving injuries to the wrist which occurred at coal or metal/nonmetal mines during the period 1990-1993.

The results of this search are shown in figure 10. There were 610 accidents that matched the search criteria. Each accident appears as a row in the table. For each accident, the three requested Search Output fields are reported as columns of the table. For example, the second record identifies an incident that was reported in 1993, occurred on 1/08/93, and involved a mechanic/repairman. REGULAR JOB TITLE is a coded field that has been converted to the equivalent English descriptions at the user's request. To indicate this, MADSS added the word "DESC" to the field name.

This search could be easily modified to produce a list of unique companies that experienced the accidents in figure 10 by removing all Accident fields (i.e. DATE OF ACCIDENT and REGULAR JOB TITLE) from the Search Output, and then choosing one or more Company fields to report for those companies.

It is also possible to report both Company and Accident data in the search results. In that case, the search will produce a list of accidents, but each retrieved accident record will show information about the accident as well as its associated company.

Answer		
YEAR	DATE OF ACCIDENT	REGULAR JOB TITLE DESC
1993	1/06/93	NEC
1993	1/08/93	MECHANIC/REPAIRMAN
1993	1/13/93	LABORER/MUCK/MACHINE OPERATOR/PIPE GIN
1993	1/19/93	BELT/CONVEYOR MAN
1993	1/27/93	MECHANIC/REPAIRMAN
1993	2/08/93	LABORER/MUCK MACHINE OPERATOR/PIPE GIN
1993	2/26/93	HIGHLIFT OPERATOR/FRONTEND LOADER
1993	3/03/93	HIGHLIFT OPERATOR/FRONTEND LOADER
1993	2/26/93	SHUTTLE CAR OPERATOR/RAM CAR
1993	1/14/93	LABORER/MUCK MACHINE OPERATOR/PIPE GIN
1993	2/18/93	MECHANIC/REPAIRMAN
1993	1/21/93	BULLDOZER OPERATOR/TRACTOR/HEAVY EQUIPMENT
1993	3/15/93	MECHANIC/REPAIRMAN
1993	3/19/93	MECHANIC/REPAIRMAN
1993	4/20/93	TIMEKEEPER/CLERK/OFFICE HELP
1993	4/12/93	SCOOP CAR OPERATOR/UNITRAC OPERATOR
1993	4/29/93	LABORER/MUCK MACHINE OPERATOR/PIPE GIN

1 of 610

Figure 10: Output for the ungrouped search

GROUPING THE SEARCH OUTPUT

By default, the search in figure 9 extracted an itemized list of accidents that matched the search criteria. However, MADSS provides an optional Grouping feature that can be used to group the search results according to user-defined criteria and calculate statistics for those groupings. This feature is very useful for identifying accident trends.

For example, figure 11 shows the accidents in figure 10 grouped by YEAR, with a "count" statistic computed for DATE OF ACCIDENT. The count statistic computes the number of nonblank entries found for the selected field in each category. Because DATE OF ACCIDENT is never blank, this effectively computes the number of accidents in each grouping. According to figure 11, 114 of the accidents occurred in 1990, 132 occurred in 1991, 165 occurred in 1992, and 199 occurred in 1993 (note that the sum of these accident counts equals the original 610 accidents that were extracted in the ungrouped search). These results appear to indicate an increase in repeated trauma, wrist injuries reported over this four year period (although one would want to normalize the rates based on actual work hours to be certain).

C:\OUTPUT	Answer	
	YEAR	CNTDATEOFACCIDENT
1	1990	114.00
2	1991	132.00
3	1992	165.00
4	1993	199.00

1 of 4

Figure 11: Output for the grouped search

A maximum of five statistics may be calculated per field. Other valid statistics include minimum, maximum, sum, mean, variance, population standard deviation, and sample standard deviation. For example, one could compute the mean victim job experience for each category.

By carefully selecting the Search Output fields and statistics, extracted data can be grouped in a variety of ways. For example, one could compute the number of machinery accidents grouped by state of location, equipment make and model, or victim job title. Or, accidents could be grouped by degree of injury (fatality, lost time, etc.), with mean victim age calculated for each category. Grouping categories may consist of more than one field. For example, a list of accidents could be grouped by both state of location *and* degree of injury. In this case, a count computed for DATE OF ACCIDENT would identify the number of fatalities, lost time accidents, etc. that occurred in each state for each year.

MISCELLANEOUS FEATURES

MADSS provides various utilities for working with data retrieved in a search. The View utility lets the user calculate statistics on, or locate particular values within, a column of data. For example, this feature could be used to find all occurrences of the word "operator" in the REGULAR JOB TITLE column in figure 10. The Sort utility lets the user sort the extracted data into a different order. For example, the table of results in figure 10 could be sorted by REGULAR JOB TITLE, in either ascending or descending order. The Report utility can generate a printable report of the search results. Finally, the Export utility lets the user reformat the extracted data for use with other software packages. Therefore, users can easily translate data to a format that can be recognized by their favorite spreadsheet, database, word processing, graphing, or slide generation software.

SUMMARY

The sample searches in this paper demonstrate just a few ways that MADSS can be used to search the MSHA Company and Accident databases. Because MADSS lets the user control what information to extract and how to present it, it can accommodate the data analysis requirements of users with a variety of backgrounds and interests. For example, companies can use MADSS to generate a list of their own accidents, or to compare their accident rates to companies with similar characteristics. State agencies and MSHA inspection offices can use the software to analyze data for their geographical areas of interest. Researchers with particular focus areas, such as back injuries or pneumoconiosis, will also find that MADSS can meet their needs.

The software provides many features for searching the MSHA data. The Grouping option is especially useful for identifying accident trends, although it is also possible to generate itemized lists of companies or accidents.

Online help is available for working with the MSHA data, and MADSS can translate coded MSHA data to their equivalent English descriptions. This eliminates the need to refer to paper manuals when interpreting the search results.

An Export feature is also available to reformat search output for use with third party software, such as dBASE, Lotus, and Quattro Pro. This makes it easy to transfer data to other software applications for advanced statistical analysis, report formatting, or graph generation.

MADSS will be completed in mid-1996. At that time, the software and revised MSHA data files for 1978 through 1995 will be made available to the public for downloading via the World Wide Web (WWW) and FTP. Although the software is a standalone PC package, the MSHA data can be installed on a network file server for shared access by multiple users. A CD-ROM version is also anticipated. The CD's will be updated on an annual basis, and will include data for completed years. Future data, including quarterly data updates, will be posted on the Internet. CD users will also be able to download the quarterly updates (to be installed to a supplementary local or network hard drive) for use until a revised CD is available.